

# Scale-Space Signatures for the Detection of Clustered Microcalcifications in Digital Mammograms

Thomas Netsch\* and Heinz-Otto Peitgen

**Abstract**—A method is described for the automated detection of microcalcifications in digitized mammograms. The method is based on the Laplacian scale-space representation of the mammogram only. First, possible locations of microcalcifications are identified as local maxima in the filtered image on a range of scales. For each finding, the size and local contrast is estimated, based on the Laplacian response denoted as the scale-space signature. A finding is marked as a microcalcification if the estimated contrast is larger than a predefined threshold which depends on the size of the finding. It is shown that the signature has a characteristic peak, revealing the corresponding image features. This peak can be robustly determined. The basic method is significantly improved by consideration of the statistical variation of the estimated contrast, which is the result of the complex noise characteristic of the mammograms. The method is evaluated with the Nijmegen database and compared to other methods using these mammograms. Results are presented as the free-response receiver operating characteristic (FROC) performance. At a rate of one false positive cluster per image the method reaches a sensitivity of 0.84, which is comparable to the best results achieved so far.

**Index Terms**— Clustered microcalcifications, image noise, Laplacian scale-space, mammography.

## 1. INTRODUCTION

**B**REAST cancer is the most common malignant disease among women in the USA and Europe. Statistics indicate that in the USA approximately one in ten women will develop breast cancer during lifetime [1]. Mammography is widely regarded as the most effective method for the early detection of breast cancer, since tumors may be visualized by indirect signs such as the presence of clustered microcalcifications and architectural distortion. The radiological interpretation of mammograms, however, is a difficult task, since the mammographic appearance of normal tissue is highly variable. The nature of many abnormal patterns is subtle and their perception is often limited by low attenuation contrast.

In particular, the effectiveness of breast cancer screening, where abnormalities have to be detected at an early stage in a large number of asymptomatic women, depends on the radiologist's ability to detect signs of malignancy. For this reason, independent reading of screening mammograms by two

radiologists is suggested to reduce the number of interpretation errors. In spite of this, retrospective analysis of previous mammograms still shows a substantial fraction of missed cancers. Considerable research has been undertaken in the development of automated image analysis methods to assist radiologists in the identification of abnormal lesions. Recent studies demonstrate that computer-based prompting systems, which indicate suspicious regions in the mammogram, have promising potential to reduce the number of lesions missed by radiologists [2], [3].

Much of the current work on computerized detection of mammographic abnormalities focuses on the recognition of clustered microcalcifications [4]–[7]. This might be partly due to their clinical importance, about 30–50% of mammographically detected cancer contains microcalcifications visible in the image [8], but is also due to their characteristic image features, which can be addressed by various pattern recognition techniques.

Microcalcifications are calcium deposits as small as 0.1–0.3 mm in diameter which can be identified in mammograms as tiny areas that are slightly brighter than their surrounding tissue. Only clusters of three or more particles within a region of approximately 1 cm<sup>2</sup> are considered as clinically suspicious [9]. The morphologic appearance of microcalcifications has been emphasized in the last decades [10]. Pleomorphic, curvilinear, or branched shapes are usually found within malignant lesions, while sharply outlined pearl-like appearances are associated with benign growth processes. A mammogram containing a malignant-type microcalcification cluster is shown in Figs. 1 and 2.

Detection algorithms are frequently evaluated with a local database, which makes it difficult to objectively compare the performance of the different methods. We are using a publicly available database,<sup>1</sup> the Nijmegen database [11], which has been widely used by researchers for performance measurements. In the following, we briefly discuss previous work evaluated with this database. Kegelmeyer and Allmen [12], for example, apply several image segmentation methods and image features to identify microcalcifications. The features are classified by binary decision trees. The features determined for training images are balanced with truth information to grow the trees. Pixels of unknown mammograms are then labeled, due to the trees yielding a map indicating the probability that a pixel belongs to a microcalcification. The performance of their approach, however, is rather poor, which they partly explain by

Manuscript received October 13, 1998; revised April 26, 1999. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was A. Burgess. Asterisk indicates corresponding author.

\*T. Netsch is with the Philips Research Laboratories Hamburg, Roentgenstrasse 24-26, D-22335 Hamburg, Germany.

H.-O. Peitgen is with the Center for Medical Diagnostic Systems and Visualization (MeVis), D-28359 Bremen, Germany.  
Publisher Item Identifier S 0278-0062(99)09059-X.

<sup>1</sup><http://marathon.csee.usf.edu/Mammography>

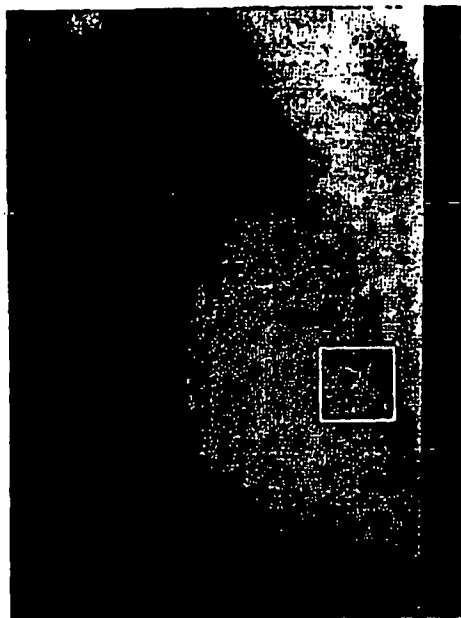


Fig. 1. Nijmegen mammogram c050 consisting of a microcalcification cluster in the marked region (see also Fig. 2). The mammogram is a mediolateral oblique view of the right breast.



Fig. 2. Microcalcification cluster marked in the mammogram shown in Fig. 1. The size of the region is about 2 cm (200 pixel).

the use of impure training data. Meersman *et al.* [13] use neural networks with varying dimensions. Each pixel in a  $64 \times 64$  region is trained by feeding its squared neighborhood of  $9 \times 9$  pixels as the input to the network. The performance of the network is optimized by means of the mean square error of the network output and the correct segmentation. Kimmig *et al.* [14] apply matched filter theory for the detection and area and shape features to eliminate various false detections. A wavelet approach for the recognition of individual microcalcifications is suggested by Strickland and Hahn [15]. They implemented a multiscale matched filter bank by means of a suitable wavelet transform, which is superior to the results of Kimmig. A method based on adaptive filtering and a Gaussianity test of the prediction error is proposed by Gürcan *et al.* [16]. For the prediction of pixel values, a two-dimensional (2-D) least mean square adaptive filter is applied. Using experimentally determined parameters, the scheme is able to detect all true

microcalcification clusters of the Nijmegen database, with 2.3 false clusters per image although a different overlap criterion is used for the evaluation. The best results for the database obtained so far, however, are by Karssemeijer [11]. He describes a Bayesian method utilizing a Markov random field (MRF) model [17], [18]. We will later focus on this approach in more detail.

The paper is organized as follows. In Section III, a novel method for the detection of individual microcalcifications is introduced [19], [20]. Potential microcalcifications are first identified with a multiscale spot detector based on Laplacian filter kernels. Then, their size and local image contrast are robustly estimated, using the filter response of their center pixel. The basic method is further improved by consideration of the statistical variation of the extracted features caused by image noise. In Section IV, the method is evaluated with the Nijmegen database and the improvements are demonstrated. Finally, in Section V we compare the results with Karssemeijer's MRF algorithm.

## II. MATERIAL

The Nijmegen database contains 40 digitized mammograms, composed of both craniocaudal and mediolateral oblique views recorded from 21 patients. The images are digitized with a CCD camera from screen-film at a pixel size of 0.1 mm and a 12-bit grayscale linear to the transmitted light during digitization. The image matrix is  $2048 \times 2048$  pixels. Each mammogram shows one or more microcalcification clusters verified by histology. The total number of clusters in the database is 105. The location and size of each cluster is given by a truth circle marked by expert radiologists. The database also contains look-up tables for rescaling of the image data and text files, storing the center and diameter of the truth circles in pixel coordinates.

The evaluation of the performance of computerized detection methods is usually addressed by free-response receiver operating characteristic (FROC) methodology [21], [22]. FROC analysis allows us to take the location of abnormal lesions in the mammogram into account. Images with multiple clusters and multiple computer-generated prompts can therefore be used within the evaluation. Based on the ground truth information provided by the database the number of true positive (TP) and false positive (FP) clusters are counted according to a given level of detection sensitivity, which is controlled by a corresponding parameter setting of the algorithm. The FROC diagram represents the fraction of TP clusters (TP fraction) as a function of the number of the FP clusters per image. Each parameter setting results in a corresponding operating point in the diagram. Several evaluations are needed to describe the performance of the method. The assessment of the significance of the results is usually addressed by fitting the experimental data to an underlying statistical model of observer performance. Although quantitative measures have been proposed for FROC analysis to compare the results of different detection schemes [22], such measures are rarely used in work on automated detection for significance testing.

FROC results depend on the applied test, which defines whether the computer findings are regarded as FP or TP

clusters. To effectively compare the performance of different detection methods, not only the same set of images must be used, but also the same overlap rule. Following Karssemeijer [11], a 1 cm (100 pixel) diameter circle is inscribed around each individual finding. Intersecting disks are merged to a single connected region. A true cluster is considered to be detected if at least two findings are found in the associated truth circle marked by expert radiologists. Regions without findings in truth circles are counted as FP clusters if they contain more than one finding. The overlap rule also removes isolated findings within the 1 cm circle. Such findings are frequently caused by film scratches and emulsion errors.

### III. METHOD

#### A. Motivation and Outline

At a spatial resolution of 0.1 mm pixel size, microcalcifications cover only a small number of image pixels and appear as small bright circular or slightly elongated spots in the mammogram. Therefore, we plan to solve the detection problem as follows:

- 1) find bright, almost circular spots in the mammogram;
- 2) estimate size  $D$  and local contrast  $C$  of each spot;
- 3) mark a spot as a microcalcification if  $C > C_T(D)$  where  $C_T$  is a given threshold depending on the estimated size  $D$  of the spot.

The threshold  $C_T$  is used to control the sensitivity of the detection. For example, if  $C_T$  is small and also diffuse, faint spots in the mammogram are regarded as possible microcalcifications. The threshold  $C_T$  depends on the size  $D$  since large solid microcalcifications absorb more X-ray energy and appear brighter in the mammogram.

Spots of a different size are detected in step 1) by a method proposed by Blostein and Ahuja [23]. The authors show that bright spots correspond to local maxima in Laplacian convolved images if the size of the filter kernel is chosen appropriately. To reliably detect all potential microcalcifications, local maxima must be considered on a range of scales. Methods concerning the scale of objects to be detected by using convolutions for different kernel sizes are referred to as scale-space approaches [24], [25].

Further motivation for the application of Laplacian kernels can be found in the context of matched filter theory [26]. A matched filter maximizes its response at the location of a well-defined known type of pattern in a noisy image. Detection is accomplished by sampling the output of the filter for values exceeding a chosen threshold. Strickland and Hahn [15] show that the high-frequency noise of mammograms can be modeled by a separable Markov process and that the average gray level profile of microcalcifications is well described by a circularly-symmetric Gaussian function. With these assumptions, the resulting matched filter is a Laplacian kernel.

For the estimation of the size in step 2) we initially applied standard methods, such as region growing techniques. These algorithms, however, depended on many implicit parameters and were not very stable since assumptions on the local contrast of the marked spots had to be made. Due to these

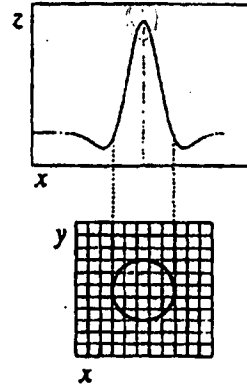


Fig. 3. Cross section of the 2-D Laplacian filter with the  $xx$  plane for scale  $h = 5$  and the corresponding view on the  $xy$  plane. The circle indicates the intersection of the lobe with the plane.

difficulties, we developed a method based on a model of a microcalcification and its properties in the scale-space. We will show that the Laplacian response of a spot, its scale-space signature, reveals the underlying parameters of the corresponding model. Finally, in step 3), we choose for each size  $D$  a large number of different threshold values to evaluate the FROC performance of the method.

#### B. Laplacian Scale-Space

The 2-D Laplacian-of-Gaussian (LoG) is defined by the application of the Laplacian operator  $\Delta$  to the Gaussian

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (1)$$

where  $\sigma > 0$  is the standard deviation and  $x, y \in \mathbb{R}$ . Additionally, a normalization factor  $-n(\sigma)$  is introduced yielding

$$\begin{aligned} \text{LoG}_\sigma(x, y) &= -n(\sigma)(\Delta G_\sigma(x, y)) \\ &= n(\sigma) \frac{1}{\pi\sigma^4} \left(1 - \frac{x^2+y^2}{2\sigma^2}\right) e^{-(x^2+y^2)/2\sigma^2}. \end{aligned} \quad (2)$$

The continuous parameter  $\sigma$  is denoted as the scale of the filter. An important question is, for which values of  $\sigma$  the image has to be convolved. Since images are given as a matrix of pixels, it is straightforward to measure the scale  $\sigma$  as a number  $h$  of pixels. As depicted in Fig. 3, the scale of the filter is defined as the diameter of the central lobe, the positive part of the function above the  $xy$  plane, measured in pixels. The lobe intersects the  $xy$  plane in a circle with diameter  $2\sqrt{2}\sigma$ , thus, for a given scale  $h$ , the corresponding value of  $\sigma$  is

$$\sigma = \frac{1}{2\sqrt{2}} h. \quad (3)$$

We define the scale-space representation of an image  $f$  as the convolution of the Laplacian filter and the image

$$[\text{LoG}_h * f](x, y) \quad (4)$$

for the scales  $h = 1, \dots, h_{\max}$ .

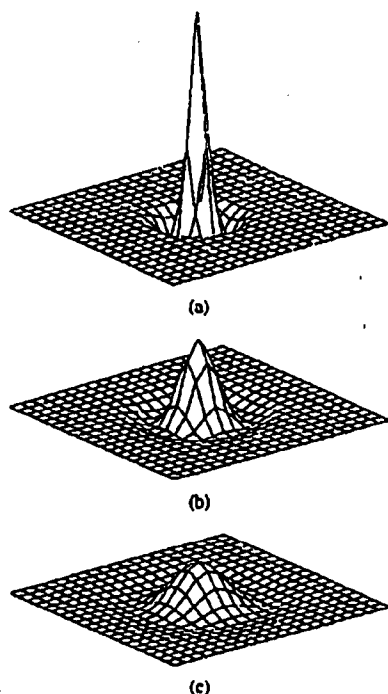


Fig. 4. Examples of normalized Laplacian filter kernels for scales (a)  $h = 3$ , (b)  $h = 5$ , and (c)  $h = 7$ .

The volume of the central lobe, calculated to  $2/e\sigma^2$ , depends on  $\sigma$ , which is inconvenient if Laplacian responses on different scales are directly compared to each other. Therefore, the filter is normalized with respect to the volume of the lobe. Hence, with (3) it follows

$$n(h) = \frac{1}{16} e h^2. \quad (5)$$

Fig. 4 shows examples of normalized Laplacian filter kernels.

The Laplacian must be carefully sampled, since only a few values approximate the function for small scales. Following Neyenssac [27], the size of the discrete filter kernel is chosen to be  $3h$ . Even kernel sizes are increased by one. The fundamental properties of the Laplacian, the integral with respect to the  $xy$  plane is zero and the integral with respect to the central lobe is one due to the normalization, should also hold for discrete kernels. The constraints require an additional factor  $k_1 > 0$  and a relaxation parameter  $k_2 \geq 1$  in (2)

$$\text{LoG}_\sigma(x, y) = k_1 n(\sigma) \frac{1}{\pi\sigma^4} \left( 1 - k_2 \frac{x^2 + y^2}{2\sigma^2} \right) \cdot e^{-(x^2 + y^2)/2\sigma^2}. \quad (6)$$

The values of  $k_1$  and  $k_2$  are determined such that both constraints hold for the discrete kernels. For  $h = 1$  the values are  $k_1 = 3.461$  and  $k_2 = 3.537$ , for  $h = 2$  the values are  $k_1 = 1.156$  and  $k_2 = 1.002$ , and for  $h \geq 3$  we have  $k_1 = 1.000$  and  $k_2 = 1.000$ , respectively. Because of the large relaxation values, scale  $h = 1$  is not considered. Moreover, the

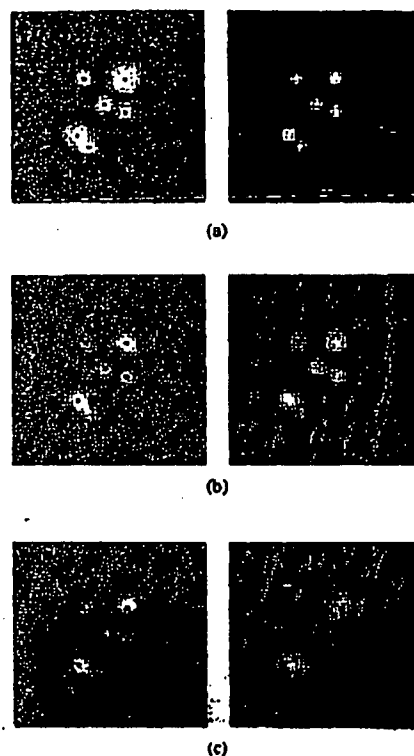


Fig. 5. Artificial image consisting of bright spots of different size and local contrast. As well, Gaussian noise with zero mean and fixed standard deviation is added. In the left row the black pixels mark the location of local maxima of the Laplacian filtered image for scales (a)  $h = 2$ , (b)  $h = 5$ , and (c)  $h = 10$ . The corresponding filtered images are shown on the right.

filter kernels are separated in one-dimensional (1-D) kernels to reduce the computational complexity of the convolution.

### C. Spot Detection

Possible microcalcifications correspond to local maxima in the Laplacian filtered image. A pixel is considered as a local maximum if its value is the largest in a  $3 \times 3$  neighborhood. Due to the smoothing properties of the Gaussian, a  $3 \times 3$  neighborhood is sufficient to robustly identify local extrema for large filter kernels, as well. The local maxima are denoted as candidate pixels in the following.

To understand how candidate pixels are generated by the Laplacian, an artificial image is investigated. The image consists of several bright spots of different sizes and local contrast, as shown in Fig. 5, left row. Moreover, Gaussian noise with zero mean and constant standard deviation is added. The image and the Laplacian kernel are convolved on scales  $h = 2$ ,  $h = 5$ , and  $h = 10$ . On scale  $h = 2$  several artifacts are generated in the background area, due to the image noise. These pixels mark local maxima with very small local contrast. For larger scales, the Gaussian blurring becomes apparent: the image noise and smaller spots are removed while candidate pixels of larger spots are dislocated. On scale  $h = 10$ , the

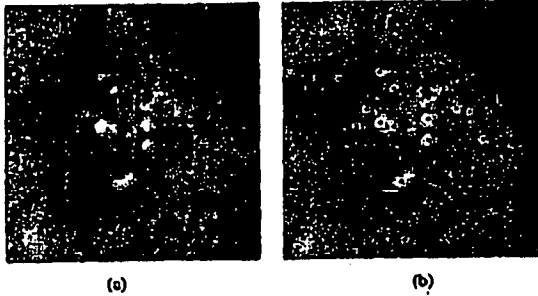


Fig. 6. (a) Patch of mammogram c010 consisting of a microcalcification cluster. (b) The candidates are marked by the Laplacian detection applying path tracking as black dots. The size of the patch is 1.5 cm (100 pixels correspond to 1 cm).

influence of image structures in the neighborhood is evident: some of the local maxima are shifted from their true center.

For reliable estimates of the size and local contrast of a bright spot, which is explained in the next section in more detail, it is important that spots are marked at their true center. Therefore, the localization of corresponding candidate pixels is improved by coarse-to-fine tracking, which is a well-known technique in applications of scale-space filtering [28]–[30]. For example, large-scale edges in an image can be identified on a coarse scale, but more accurately located on a finer scale by following the zero crossings of the Laplacian. Here, displacements are identified by tracking the position of the local maxima in the scale-space. A path always starts with a candidate pixel at the finest scale  $h = 2$  and ends at some coarser scale if an appropriate candidate pixel cannot be found in the intermediate  $3 \times 3$  neighborhood. The length of a path in the scale-space depends on the local structure of the mammogram. Long paths rarely occur and most paths already end at scale  $h = 4$ .

The candidate pixels of a path belong to the same spot, but their correlation with the true center of the spot may vary. Due to the normalization of the filter kernel with respect to the central lobe, the magnitude of the response is an indicator of the correspondence between the position of the candidate pixel and the actual center pixel of the spot. The absolute maximum of the response along each path indicates the optimal candidate representing the spot in the image which is reported for each path. Fig. 6 depicts a patch of mammogram c010 and the results of the Laplacian detection with path tracking.

#### D. Feature Estimation by Signatures

The estimation of the local contrast and the size of a spot is based on the scale-space representation of the image. First, a microcalcification is modeled by a cylinder of a specific height  $C$  and diameter,  $D$ , which corresponds to the local contrast and the diameter of the spot in the image, respectively. An image  $f$ , consisting of a single cylinder with its center at  $M = (0,0)$ , is defined by

$$f(x, y) = \begin{cases} C, & \text{if } x^2 + y^2 \leq \frac{D^2}{4} \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

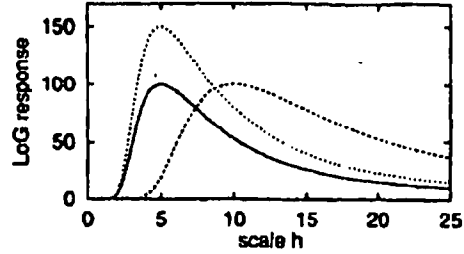


Fig. 7. Signature of the cylindrical microcalcification model. The solid curve shows the Laplacian response of the model in  $M$  for  $C = 100$  and  $D = 5$ . A change to  $C = 150$  (higher peak) or  $D = 10$  (peak to the right) yield the dashed curves, respectively.

It is not necessary to introduce an additional parameter for the image background because Laplacian convolution is independent of a constant background. Obviously, for a fixed scale the center,  $M$  is a local maximum in the Laplacian filtered image.

The response  $r_{cyl}(\sigma)$  in  $M$  is analytically calculated by

$$r_{cyl}(\sigma) = [\text{LoG}_\sigma * f](x, y)|_{M=(0,0)}. \quad (8)$$

With (2) and (7) and polar coordinates  $u = r \cos(\phi)$  and  $v = r \sin(\phi)$ , it follows that

$$\begin{aligned} r_{cyl}(\sigma) &= n(\sigma) \frac{C}{\pi \sigma^4} \int_0^{D/2} \int_0^{2\pi} \left(1 - \frac{r^2}{2\sigma^2}\right) \\ &\quad \cdot e^{-(r^2/2\sigma^2)} r d\phi dr \\ &= n(\sigma) \frac{2C}{\sigma^4} \int_0^{D/2} \left(1 - \frac{r^2}{2\sigma^2}\right) e^{-(r^2/2\sigma^2)} r dr. \end{aligned} \quad (9)$$

The remaining integral can be calculated to  $(D^2/8)e^{-(D^2/8\sigma^2)}$  and with (3) and (5) it is

$$r_{cyl}(h) = C e^{-\frac{D^2}{h^2}} e^{-(D^2/h^2)}. \quad (10)$$

The graph of  $r_{cyl}(h)$ , shown in Fig. 7, is referred to as the signature of the model. The signature has a characteristic peak which reflects the underlying parameters of the model. By differentiation of (10), the coordinate  $h_{max}$  of the local maximum in the graph evaluates to

$$h_{max} = D \quad (11)$$

and

$$r_{cyl}(h_{max}) = C. \quad (12)$$

A candidate pixel in the mammogram has a reasonable correlation with the microcalcification model. Therefore, the size  $D^*$  and the local contrast  $C^*$  of the spot are determined from the actual signature  $r^*(h)$  by searching for the position  $h_{max}$  of the characteristic peak. With (11) and (12) it follows  $D^* = h_{max}$  and  $C^* = r^*(h_{max})$ . A visual inspection shows that candidate pixels in the mammogram have a well-defined peak, which results in a stable and robust estimation.

In Section III-C, path tracking was introduced to guarantee that the artificial spots in Fig. 5 are marked at their true center, since the proposed estimation is only valid for signatures calculated in  $M$ . In the case of mammograms, surrounding bright

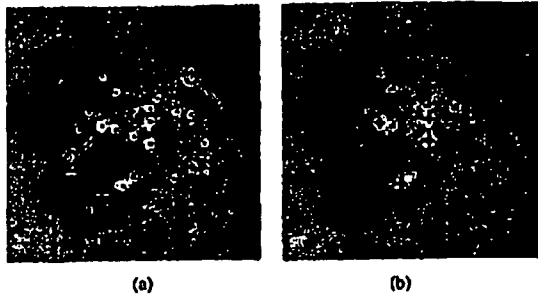


Fig. 8. (a) Patch of mammogram c010 shown in Fig. 6. The signature of three candidate pixels (marked by 1, 2, and 3) is shown in Fig. 9. (b) A finding reported as microcalcification is marked with a cross. The size of a cross corresponds to the estimated size.

tissue might displace candidate pixels as well, in particular, for larger filter kernels. Therefore, path tracking is applied to improve the correlation with the model and to accurately estimate the local contrast of candidates.

In Fig. 9, signatures for several candidate pixels in the mammogram are given. For example, the first candidate has its local maximum at scale  $h_{\max} = 6$  resulting in an estimated size of  $D^* = 6$ . The contrast  $C^*$  is larger than the predefined threshold  $C_T(D^*)$  and the candidate is marked as a microcalcification. The response of the second candidate has two local maxima for  $h_{\max} = 3$  and  $h_{\max} = 14$ . The maximum on the coarser scale, however, is not taken into account, since image structures in the neighborhood dominate the signature. As in the first example, this candidate is also reported as a microcalcification. In the last example, we have a clear peak at  $h_{\max} = 3$ , but the estimated local contrast  $C^*$  is smaller than  $C_T(D^*)$  and the candidate is not marked. Fig. 8 also depicts the results for the microcalcification cluster already shown in Fig. 6. It should be noted that from the detection point of view, there is no need to segment individual microcalcifications from the background tissue.

Following Strickland and Hahn [15], the 2-D Gaussian

$$f(x, y) = C e^{-(4/D^2)(x^2 + y^2)} \quad (13)$$

is a more reasonable model for a microcalcification. Similar to the derivation of (10) the response  $r_{\text{gauss}}(h)$  in  $M = (0, 0)$  is calculated as

$$r_{\text{gauss}}(h) = C e^{\frac{D^2 h^2}{(D^2 + h^2)^2}}. \quad (14)$$

The position of the local maximum evaluates to

$$h_{\max} = D \quad (15)$$

and

$$r_{\text{gauss}}(h_{\max}) = \frac{C}{4}. \quad (16)$$

Due to these equations, the cylindrical and the Gaussian model are equivalent if the response  $r_{\text{gauss}}(h)$  is properly scaled. In

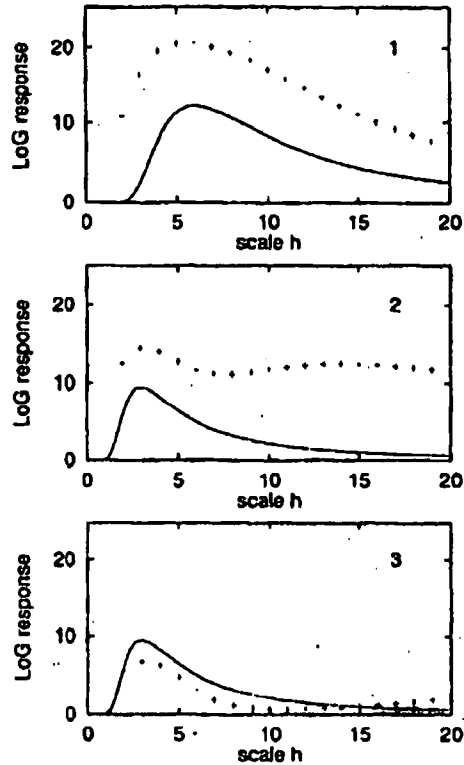


Fig. 9. Signature (+) of three candidate pixels, shown in Fig. 8(a), compared to the signature of the model for the estimated size  $D^*$  and a predefined contrast threshold  $C_T(D^*)$  (solid).

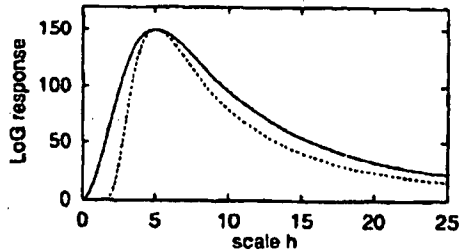


Fig. 10. Signature of the Gaussian microcalcification model. The solid curve shows the Laplacian response of the model in  $M$  for  $C = 220$  and  $D = 5$ . The signature is compared to the response of the cylindrical model with  $D = 5$  and the corrected contrast parameter.

Fig. 10, the graph of  $r_{\text{gauss}}(h)$  is compared to the response of the cylindrical model. The algorithm could only take advantage of the slight difference between the curves near their local maximum if the comparison of  $C^*$  and  $C_T(D^*)$  is extended to a small neighborhood around  $D^*$ . The same arguments also hold for elongated ellipsoidal models, which might be more suited for the malignant type of microcalcifications. A visual inspection of the candidate signatures in the mammograms shows that the response near the characteristic peak is often disturbed by neighboring structures. Therefore, we use only the

local contrast information at position  $D^*$ , which is sufficiently described by the cylindrical model.

### E. Image Noise

The noise characteristic of digitized mammograms is complex and strongly depends on the gray level itself. The noise level is considerably higher in bright regions of the mammogram representing glandular tissue. As a result, the statistical variation of features extracted from different regions depends on the predominant type of tissue. For example, a given threshold  $C_T$  might identify microcalcifications in dark regions, but will generate many FP detections in bright regions. If the threshold is increased the FP detections are reduced, but faint microcalcifications are lost in the darker areas of the mammogram. To guarantee the same detectability for microcalcifications in different image regions, the feature estimation algorithm must consider the gray level dependency of the noise. A solution to this problem is the use of adaptive methods which vary  $C_T$  according to the local noise variance. In work on microcalcification, detection adaptive methods can be found in Chan *et al.* [31] and Davies and Dance [4].

Karssemeijer and van Erming [11], [32] describe a different approach. They propose a noise equalization procedure applied prior to any subsequent image processing. The equalization can be understood as a nonlinear transformation  $L(y)$  of the gray values  $y$  to a scale on which the noise level is constant. The image noise is measured by the local image contrast

$$c_i = y_i - \frac{1}{N} \sum_{j \in \theta_i} y_j \quad (17)$$

where  $y_i$  is the gray level at site  $i$ ,  $\theta_i$  denotes neighboring sites around  $i$ , and  $N$  the total number of neighbors. The statistics of  $c_i$ , in particular the mean  $\mu$  and the standard deviation  $\sigma$ , are used to represent the noise characteristic.

Given a gray level band  $b = [y - k, y + k]$  with  $k \geq 0$ , a suitable sample of sites  $x$  with  $f(x) \in b$  is needed to get reliable estimates of the distribution. In mammography this is not a particular problem because of the large size of the image matrix. Karssemeijer divides the entire grayscale into a small number of predefined bands where the noise characteristic is calculated. We are using a different approach, which results in a larger number of bands and more detailed measurements. For an arbitrary gray value  $y$ , a corresponding band  $b$  is constructed by successively adding the neighboring levels  $y \pm 1, y \pm 2, \dots$  until the band consists of a minimum number of pixel sites. For each band  $b$ , the histogram of the local contrast  $c_i$  is determined, followed by the calculation of the mean  $\mu(y)$  and the standard deviation  $\sigma(y)$  of the bands.

Experiments show that already about 1500 image sites are sufficient to yield stable estimates of the moments of the histogram. The normalized histograms resemble a Gaussian distribution with zero mean. Fig. 11 depicts the estimated standard deviation  $\sigma(y)$  of the local contrast  $c_i$  as a function of the gray value for the Nijmegen mammogram c15c.

Given constant noise level  $\sigma_r$ , the scale transformation  $L(y)$  can be determined from the standard deviation  $\sigma(y)$  by numerical integration. The transformation itself is represented

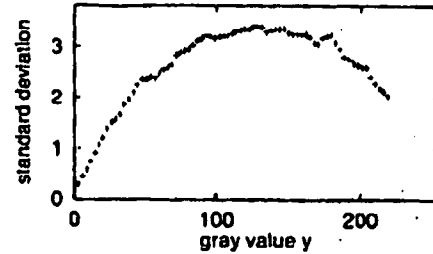


Fig. 11. Standard deviation of the image noise as a function of the gray value for the Nijmegen mammograms c15c. The original images are rescaled to 8 bits prior to the measurement. Each mark in the plot represents the result of a single band.

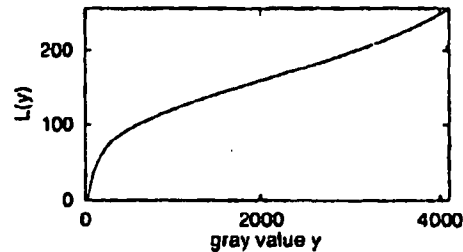


Fig. 12. Scale transformation  $L(y)$  for mammograms c15c, as provided by the Nijmegen database. The transformation also reduces the original grayscale from 12 to 8 bits.

by a look-up table. We do not intend to describe the calculation of  $L(y)$  in more detail, since the Nijmegen database already contains an appropriate look-up table for each mammogram, which we use for the experiments in this paper. However, for the calculation of the Laplacian feature noise, which is described in Section III-E2, we apply the noise estimation, as proposed above. The average noise level of the rescaled Nijmegen mammograms is  $\sigma_r = 1.98$ . In Fig. 12, the transformation  $L(y)$  for mammogram c15c is shown. In the next section, we consider consequences of the consideration of Gaussian noise.

1) *Noise-Corrected Microcalcification model:* Due to noise-equalization, mammograms have a constant noise level  $\sigma_r$ . After Laplacian convolution the noise still has a Gaussian characteristic, but the remaining noise level depends on the degree of smoothing and, hence, on the size of the filter kernel. The response of a candidate pixel should consider the noise level as a function of scale, otherwise, the local contrast is not correctly estimated.

Therefore, we complement the cylindrical microcalcification model by Gaussian noise with standard deviation  $\sigma_r$  and zero mean. Since convolution is a linear operation, signal and additive noise can be considered separately. The noise component yields the additional term  $\sigma(h)$  in the signature  $\tau_{cyl}(h)$  in (10) of the microcalcification model. For the calculation of  $\sigma(h)$ , the change of the gray value at site  $i$  is described by independent random variables  $X(i)$  with distribution  $N(0, \sigma_r)$ . The random variables  $X(i)$  are

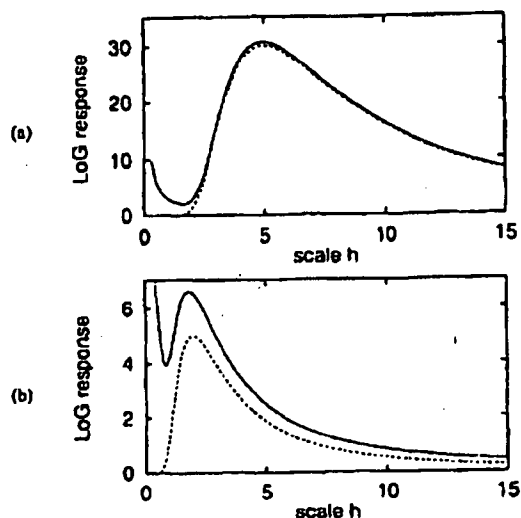


Fig. 13. Signature of the cylindrical microcalcification model with noise term (solid) and without noise term (dashed). The noise level is  $\sigma_r = 2.0$ . The local contrast and the size are (a)  $C = 30$ ,  $D = 5$  and (b)  $C = 5$ ,  $D = 2$ , respectively.

convolved with the Laplacian at scale  $h$

$$Z_h(i) = [\text{LoG}_h * Y](i) = \sum_{j \in \mathbb{Z}^2} \text{LoG}_h(j) X(i+j) \quad (18)$$

yielding new random variables  $Z_h(i)$ . Since  $Z_h(i)$  is a linear combination of independent random variables the variance of  $Z_h(i)$  evaluates to

$$\begin{aligned} \text{Var}(Z_h(i)) &= \text{Var}(X(i)) \sum_{j \in \mathbb{Z}^2} |\text{LoG}_h(j)|^2 \\ &= \text{Var}(X(i)) \frac{c^2}{\pi} \frac{1}{h^2} \\ &= \sigma_r^2 \frac{c^2}{\pi} \frac{1}{h^2} \end{aligned} \quad (19)$$

and thus, the noise level after convolution at scale  $h$  is

$$\sigma(h) = \sigma_r \frac{c}{\sqrt{\pi}} \frac{1}{h}. \quad (20)$$

Addition of (10) and (20) yields the response  $r_{\text{noise}}(h)$  of the noise-corrected cylindrical microcalcification model

$$\begin{aligned} r_{\text{noise}}(h) &= r_{\text{cyc}}(h) + \sigma(h) \\ &= C c \frac{D^2}{h^2} e^{-(D^2/h^2)} + \sigma_r \frac{c}{\sqrt{\pi}} \frac{1}{h}. \end{aligned} \quad (21)$$

The graph of  $r_{\text{noise}}(h)$  is shown in Fig. 13.

Analytic differentiation of (21) to determine the position of the local maximum is difficult. Plots of  $r_{\text{noise}}(h)$  make evident that the noise term  $\sigma(h)$  does not significantly alter the position of the characteristic peak if parameters are in the range used by the detection scheme. The model response for large values of  $C$  and noise levels  $\sigma_r$  around 2.0 is much the same as for the uncorrected model. For models with low contrast  $C$  and,

TABLE I  
NUMBER OF CANDIDATES OF MAMMOGRAM c15c FOR  
SCALES  $h = 2, \dots, 7$  IN THE CASE OF UNCORRECTED  
LAPLACIAN RESPONSES AND IN THE CASE OF NOISE CORRECTION

| rc/nh           | 2     | 3     | 4     | 5   | 6    | 7   |
|-----------------|-------|-------|-------|-----|------|-----|
| uncorrected     | 52852 | 17363 | 13080 | 815 | 3198 | 474 |
| noise-corrected | 48680 | 18101 | 15562 | 991 | 3623 | 583 |

especially, for small values of  $D$ , the noise term significantly increases the peak, which is reflected by

$$r_{\text{noise}}(D) = C + \sigma_r \frac{e}{\sqrt{\pi}} \frac{1}{D}. \quad (22)$$

Without consideration of the noise term  $\sigma(h)$ , the estimated contrast of a candidate pixel is too large, especially for faint microcalcifications with low contrast, resulting in numerous FP detections. Moreover, if mammograms with different noise level  $\sigma_r$  are considered, the sensitivity of the detection is biased by  $\sigma_r$ . Therefore, the signature of a candidate pixel is corrected by subtracting the noise term  $\sigma(h)$ . The corrected response is independent of image noise, especially of the noise level  $\sigma_r$  of a particular mammogram.

A probably more severe problem is that the image noise also affects the tracking of candidate pixels. Because of too large responses on small scales along a path in the scale-space, the absolute maximum of the responses and, hence, the optimal candidate pixel with best correlation to the model, is systematically shifted toward smaller scales. This is demonstrated in Table I. Also, the response along the path must be corrected by subtraction of the noise term given in (20). Here, the subtraction of  $1/h$  is sufficient. Constants and the image-dependent noise level  $\sigma_r$  have no effect on the search of the maximum, since only responses of the same candidate and image are compared to each other.

2) *Laplacian Feature Noise*: The decision to report a candidate pixel as a microcalcification only depends on the threshold  $C_T$  and, hence, on a reliable estimate of the local contrast of the corresponding spot, which must be independent of the actual noise characteristic. We investigate the variation of the extracted image features, which is the result of the image noise in more detail. For small scales, the local contrast kernel used for the noise estimation and the Laplacian kernel are very similar. The contrast kernel is a high-pass filter and more or less an approximation of the Laplacian. Thus, it cannot be expected that the estimation of the image noise or of the Laplacian feature noise, respectively, using either the contrast or the Laplacian convolution, are very different.

The detection algorithm, however, considers features only for candidate pixels which correspond to local maxima in the convolved mammogram. From this point of view, the variation of the Laplacian kernel should not be determined at all sites  $i$  in the mammogram, but only for the candidate sites of a given gray value band. Fig. 14 depicts the standard deviation and the mean of the Laplacian evaluated only at candidate pixels in the noise-equalized mammogram c15c. As was to be expected, the standard deviation is constant, but the mean is not zero. It shows a remarkable dependency on the gray level. If all sites are considered in the calculation, then the number of darker



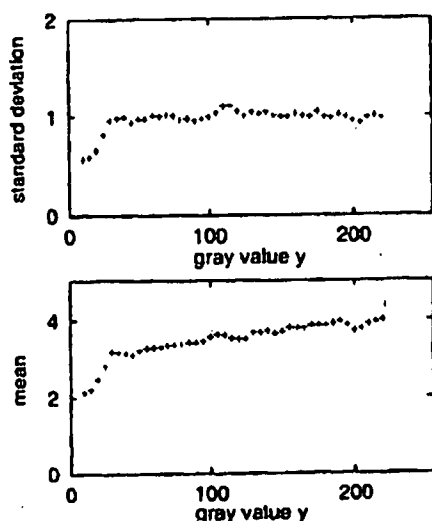


Fig. 14. Laplacian feature statistics for the noise-equalized mammogram c15c for scale  $h = 2$ . The feature is calculated only at candidate sites.

and brighter pixel values or, to be more precise, the number and latitude of positive and negative differences, is on average the same. Therefore, the mean is zero. The Laplacian is only calculated for local maxima  $i$ . The neighboring pixel values are all smaller than the value at site  $i$ , which results in a shift of the prior probabilities and thus in a nonzero mean.

The dependence of the mean on the gray level must be considered by the detection scheme as well. Otherwise, microcalcifications are not detected with the same sensitivity since the comparisons  $C > C_T(D)$  are biased by the nonconstant mean. The implementation of the adaptation is straightforward. For each mammogram and each scale, linear regression is applied to approximate the mean values by a straight line with slope  $\alpha_h$ . Let  $y$  be the gray level of a candidate pixel. Then, the actual response  $r^*(h)$  of the candidate pixel is corrected according to the slope  $\alpha_h$  by

$$r^*(h, y) = r^*(h) - \alpha_h y. \quad (23)$$

The largest slope of the line found is  $\alpha_2 = 0.0074$  (mammogram c02o), which results in a considerable change of the threshold of 1.9 for  $y = 255$ . The shift of the mean is not very pronounced for larger scales. We correct only scale  $h = 2$  and  $h = 3$ . It should be noted that the calculation of the feature statistics needs particular attention, since the local maximum constraint significantly reduces the number of sites considered in each gray level band. An analysis of the histogram for each band makes evident that a Gaussian still yields a reasonable approximation.

#### F. Implementation

Prior to any image processing the mammograms are noise-equalized by the look-up tables contained in the Nijmegen database. Each mammogram has an individual noise level  $\sigma_r$ . The breast tissue area is automatically segmented to avoid

detection artifacts caused by markers and the sharp image boundary near the chest side. To reduce the computational burden of the method, the size of microcalcifications to be detected is limited to a maximal size of  $D = 7$ . For each size  $D = 2, \dots, 7$ , a contrast threshold  $C_T(D)$  is selected to control the detection sensitivity of the algorithm. The calculation of the scale-space representation of a mammogram up to a maximal scale  $h = 12$  is sufficient to search for the local maximum of the signatures. Furthermore, all convolutions are restricted to the breast tissue area. On average, only about 39% of the image matrix has to be convolved with the Laplacian.

#### 1) Spot detection:

- find local maxima on scale  $h = 2$ ;
- initialize for each finding a path and follow the local maxima in the scale-space up to the scale  $h = 12$ ;
- correct the response along the path by subtraction of  $1/h$ ;
- identify the absolute maximum of the response and report the corresponding pixel as candidate.

#### 2) Estimation of size and local contrast:

- determine the signature  $r^*(h)$  of the candidate;
- correct the signature according to the noise term  $\sigma(h)$  given in (20);
- correct the signature according to (23) to adapt to the shift of the mean for scale  $h = 2$  and  $h = 3$ ;
- identify the characteristic peak and report  $C^*$  and  $D^*$ .

#### 3) Threshold decision:

- mark the candidate as microcalcification if  $C^*$  is larger than  $C_T(D^*)$ .

The average time needed to process a Nijmegen mammogram is about 15 s on a SGI Indigo workstation (R4400 processor, 200 MHz, 128-MByte RAM). This time is almost completely spent for the generation of the candidate pixels, which involves the convolution with the Laplacian kernels and the search for the local maxima. The corrections, the search of the peak in the signature, and the comparison with the contrast thresholds is within one second. The convolution of the entire mammographic tissue is only necessary for scales  $h = 2$  and  $h = 3$  to find the slopes  $\alpha_2$  and  $\alpha_3$ , respectively, and to identify the path entries. The feature statistics are determined by the procedure described in Section III-E. During a trace of a path in the scale-space, convolutions are only needed in a small neighborhood. Convolutions with large kernels are rare, since most paths are short. Such convolutions are implemented by a demand-driven calculation model which guarantees that each response is only determined once. The computation of signatures is also optimized. It is sufficient to determine that the signature only up to the characteristic peak is identified.

All parameters are tuned for mammograms with a 0.1 m pixel size. It turns out that a reasonable change of the parameters described above does not significantly change the detection results. Neither signatures larger than  $h = 12$  nor microcalcifications larger than  $D = 7$  improve the detection results. Also, the choice of larger neighborhoods for the search of the local maxima in the filtered mammogram, as well as for the path tracking, does not change the results.

## IV. RESULTS

The algorithm is executed for a fixed setting of thresholds  $C_T(D)$ . TP and FP clusters are counted within the mammograms of the database and the number of FP clusters per image and the relative number of TP clusters are determined according to the overlap rule. Since the threshold depends on the size  $D$ , a large number of different parameter settings must be tested to evaluate the performance of the algorithm. For example, if we use ten different thresholds for each of the sizes  $D = 2, \dots, 7$  of interest, then the entire database must be evaluated  $10^6$  times. The entire search space can be reduced to a moderate number of tests without missing settings with good results. Therefore, a set of true microcalcifications is compiled by visual inspection of several candidate pixels. Based on the histogram of the estimated contrast for each size, the search space is concentrated to a subset of 2374 different contrast settings.

To demonstrate the performance of the basic method and the improvements proposed in the previous section, in particular path tracking and the effect of image noise, we perform three FROC evaluations, summarized in Fig. 15. The plots reveal the results of all threshold settings used for the evaluation.

The basic method combines steps 1a), 1d), 2a), 2d); and 3a) of the algorithm. Since path tracking is not yet applied in the basic method, step 1a) is complemented by the local maxima for scales  $h = 3, \dots, 7$ , which introduces multiple candidate pixels for single spots. The corresponding evaluation in Fig. 15(a) shows a large number of FP detections. On average, the algorithm finds 1.5 FP clusters per image.

In the next experiment, the detection is enhanced by path tracking used in step 1b) of the algorithm. The results in Fig. 15(b) make it evident that the number of FP clusters per image significantly decreases. Compared to the basic scheme, several clusters are also missed, especially at large FP rates per image. Both observations can be explained by more accurate contrast estimates, since signatures are calculated at sites with optimal correlation to the microcalcification model because of path tracking in step 1b).

Finally, the last experiment also includes the remaining steps 1c), 2b), and 2c) which consider the effect of image noise on the signature of a candidate pixel. It can be shown that steps 2b) and 2c) contribute the same amount to the increase of the results while noise-corrected path tracking introduced in step 1c) only slightly improves the number of correctly identified clusters.

Threshold selection and performance measurement with the same set of images usually introduces a bias toward the method. To effectively compare the FROC performance of the modifications of the basic method as well as the performance of different approaches, using the Nijmegen database in Section V, the selection of optimal threshold values is addressed by a different set of 25 training mammograms, consisting of a total number of 57 microcalcification clusters. For given FP rates, settings with the best TP values are determined for the training images. If more than one setting is found for a particular FP rate, then the one with the smallest sum of thresholds on scales  $h = 2, h = 3$ , and  $h = 4$  is chosen.

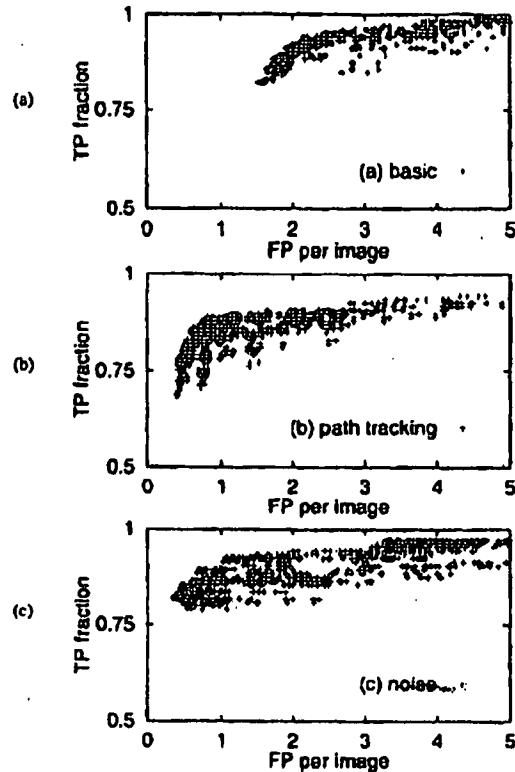


Fig. 15. (a) Evaluation of the basic algorithm. (b) Complement with path tracking and (c) noise correction. The detection results (FP fraction, number of TP clusters per image) are shown for all 2374 threshold settings with a TP rate per image smaller than 5.0.

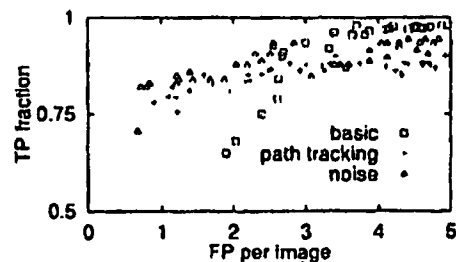


Fig. 16. Comparison of the FROC evaluation of the basic algorithm ( $\square$ ), complemented with path tracking ( $+$ ), and noise correction ( $\Delta$ ). A set of 25 training images is used to find optimal threshold settings for fixed FP rates per image.

These settings are tested against the original database yielding the FROC curves shown in the Fig. 16. Although quantitative measures are not further assessed to effectively compare the detection results, the improvement of the results due to path tracking and noise correction is apparent and comparable to the biased experiments depicted in Fig. 15.

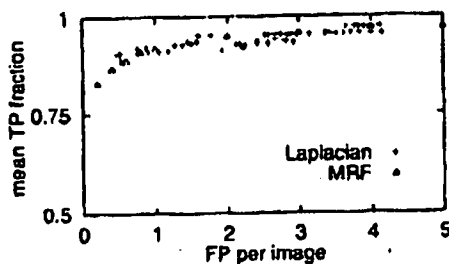


Fig. 17. Unbiased comparison of FROC evaluations of Laplacian detection (+) and the MRF approach ( $\Delta$ ). In the case of Laplacian detection, a different set of 25 training images is used to find optimal threshold settings for fixed FP rates per image.

A visual inspection of the FP clusters makes evident that, on average, all mammograms contribute the same number of FP clusters. It turns out that, especially for low threshold settings, locations on thin line structures of glandular tissue are regarded as FP clusters. The calculation of the signature for a corresponding model of a line shows a similar response as a cylindrical microcalcification model for an appropriate choice of parameters  $D$  and  $C$ . As a result, a candidate on a line might be regarded as a microcalcification for small values of  $C_T$ . An increase of  $C_T$  might eliminate line artifacts in particular cases while, on the other hand, faint calcifications may be lost.

#### V. COMPARISON

Karssemeijer's MRF approach [11] allows us to model prior knowledge of spatial relations between objects within an iterative segmentation process. As an advantage, additional information on the distribution of individual findings is included in the detection scheme. Pixels likely to be microcalcifications are rewarded if, in some spatial context, other microcalcification pixels already occur. Moreover, a line feature calculated at each image site ensures that microcalcification and line structures cannot be mixed up in an immediate neighborhood at the same time. The corresponding feature distributions are approximated by a multivariate Gaussian distribution, relying on a detailed segmentation of true microcalcifications provided by expert radiologists. The approach proposed in this paper marks individual microcalcifications. Line artifacts are only removed by the threshold criterion, introducing several FP detections, as mentioned in Section IV.

Karssemeijer uses a different nonstandard representation of the FROC results, which is based on the mean TP fraction. The mean TP fraction is defined like the usual TP fraction, but the number of TP clusters of an individual mammogram is weighted by  $1/N$  where  $N$  is the total number of true clusters in the mammogram. This guarantees that each image contributes the same amount in the FROC results, independent of the number of true clusters. The threshold settings are trained with the additional training mammograms to avoid a bias toward the Laplacian approach. As depicted in Fig. 17, the results of the Laplacian appear slightly preferable, but the MRF approach is able to operate also at FP rates smaller than 0.5 per image.

#### VI. CONCLUSIONS AND DISCUSSION

A method for the detection of individual microcalcifications based on the Laplacian scale-space representation of a mammogram is proposed. The initial identification of candidate pixels as well as the estimation of their size and local contrast uses only information provided by this representation. The sensitivity of the detection is controlled by thresholds depending on the estimated size of a candidate. The thresholds have a well-defined interpretation. A threshold value is considered as the minimal local contrast of a finding required as an indicator of a possible microcalcification. The training of the thresholds does not rely on the feature distribution or on additional estimation procedures. It is simply accomplished by a FROC evaluation of a set of training mammograms. The algorithm itself depends on very few intrinsic parameters, for example, the length of a signature or the size of the neighborhood for determination of local maxima. These parameters are not crucial with respect to the detection performance and can be adapted to mammograms with smaller pixel sizes. This results in a very stable and robust approach.

The FROC performance of the proposed method is comparable to the results of Karssemeijer's MRF approach, which achieves the best FROC results for the Nijmegen database. In particular, at 0.5 FP detections per image, the mean TP fraction is about 0.9. This is remarkable since only individual microcalcifications are addressed by the detection and the training stage depends only on additional mammograms with ground truth.

In contrast to the MRF approach, the developed algorithm is not able to operate with less than 0.3 FP detections per image at present. This can be explained by the fact that artifacts, such as scratches and film emulsion errors, are not explicitly identified and are only partly removed by the overlap rule during the FROC evaluation. Another source of FP errors is line artifacts.

An additional algorithm which recognizes findings on line structures may reduce the number of FP clusters [33]. The MRF approach certainly benefits by the additional line feature.

The FROC representation, based on the mean TP fraction chosen by Karssemeijer, might be misleading when different methods are compared. For example, the mean TP value of a cluster in a mammogram with a total number of four true clusters is only 0.25, whereas in a mammogram with just one cluster the TPF is 1.0. In the first situation four clusters have to be marked, while in the other the detection of a single cluster is sufficient to yield identical values. Table II compares the detection results for a threshold setting with an FP rate of 1.0 per image. The Laplacian detection misses only 12 clusters out of 105, compared to 17 clusters in the case of the MRF approach. This corresponds to a TP fraction of 0.89 and 0.84, respectively, but only to equal values of 0.92 in the case of a mean TP fraction representation of the FROC results.

It should be noted, that several images of the Nijmegen database, such as c12c and c17c, contain numerous overlapping clusters. It is difficult to consistently apply a cluster detection criterion in these cases, in particular, when the number of computer-generated findings is large. The counting of TP and FT clusters is somehow ambiguous then.

TABLE II  
RESULTS OF MICROCALCIFICATION DETECTION AT A FP RATE OF 1.0 PER IMAGE.  
THE TOTAL NUMBER OF TRUE CLUSTERS IS DENOTED BY  $N$ . THE  
ENTRIES OF THE TABLE IN THE FP AND FN COLUMN ARE THE RESULT  
OF THE LAPLACIAN (L) AND MRF (M) DETECTION, RESPECTIVELY

| img  | N | FP |   | FN |   | img  | N  | FP |   | FN |   |
|------|---|----|---|----|---|------|----|----|---|----|---|
|      |   | L  | M | L  | M |      |    | L  | M | L  | M |
| c01c | 3 | 0  | 1 | 0  | 1 | c11o | 1  | 1  | 2 | 0  | 0 |
| c01u | 3 | 0  | 0 | 0  | 1 | c12c | 15 | 1  | 2 | 2  | 6 |
| c02c | 2 | 0  | 1 | 0  | 0 | c12o | 13 | 1  | 2 | 2  | 2 |
| c02n | 1 | 3  | 2 | 0  | 0 | c13c | 1  | 1  | 0 | 0  | 0 |
| c03c | 2 | 0  | 0 | 1  | 1 | c13o | 1  | 0  | 0 | 0  | 0 |
| c03n | 1 | 0  | 1 | 0  | 0 | c14c | 2  | 0  | 0 | 0  | 0 |
| c04c | 2 | 1  | 0 | 0  | 0 | c14o | 2  | 1  | 0 | 1  | 1 |
| c04n | 2 | 0  | 0 | 0  | 0 | c15c | 1  | 0  | 1 | 0  | 0 |
| c05c | 1 | 1  | 1 | 0  | 0 | c15o | 1  | 1  | 1 | 0  | 0 |
| c05u | 2 | 0  | 0 | 1  | 0 | c16c | 1  | 0  | 0 | 0  | 0 |
| c06c | 3 | 1  | 0 | 0  | 1 | c16o | 1  | 0  | 1 | 0  | 0 |
| c06u | 2 | 0  | 3 | 0  | 0 | c17c | 9  | 2  | 1 | 1  | 0 |
| c07c | 1 | 2  | 0 | 0  | 0 | c17o | 5  | 1  | 2 | 0  | 1 |
| c07u | 1 | 1  | 1 | 0  | 0 | c18c | 2  | 1  | 1 | 0  | 0 |
| c08c | 4 | 0  | 0 | 1  | 1 | c18u | 1  | 0  | 0 | 0  | 0 |
| c08u | 6 | 0  | 1 | 0  | 2 | c19c | 1  | 1  | 4 | 0  | 0 |
| c09c | 1 | 0  | 1 | 0  | 0 | c19u | 2  | 1  | 0 | 0  | 0 |
| c09u | 2 | 1  | 1 | 1  | 0 | c20c | 1  | 8  | 0 | 0  | 0 |
| c10c | 1 | 0  | 1 | 0  | 0 | c21o | 1  | 5  | 3 | 0  | 0 |
| c11c | 1 | 4  | 8 | 0  | 0 |      |    |    |   |    |   |

The method described here detects individual microcalcification. Particular knowledge about the image domain, clinically suspicious microcalcifications appear in clusters, is not used so far and leaves space for improvements to the approach. The application of a neighborhood interaction model for individual findings seems most promising. A probabilistic relaxation scheme [34], a spatial scan statistic [35], or an additional clustering transformation, as suggested by Nishikawa *et al.* [36], might therefore be used.

The spatial resolution of the Nijmegen mammograms of 0.1 mm pixel size may appear rather limited compared to the actual size of faint microcalcifications. Most work on microcalcifications, however, is evaluated with mammograms of 0.1 mm pixel size. The dependence of detection accuracy on the pixel size and depth of the digitized mammograms is addressed in a study by Chan *et al.* [37]. Their results indicate that detection accuracy is essentially independent of a pixel depth from 12 bits to 9 bits, but significantly decreases as the pixel size increases. Robustness of detection algorithms against changes of the spatial resolution of images, however, might depend on the applied method as well.

#### ACKNOWLEDGMENT

The mammograms were provided courtesy of the National Expert and Training Center for Breast Cancer Screening and the Department of Radiology at the University of Nijmegen, the Netherlands. The authors are gratefully indebted to N. Karssemeijer for providing the 25 additional training mammograms.

#### REFERENCES

- [1] M. J. Broeders and A. L. Verbeek, "Breast cancer epidemiology and risk factors," in *Radiological Diagnosis of Breast Diseases*, M. Friedrich and E. A. Sickles, Eds., Berlin, Germany: Springer-Verlag, 1997, pp. 1-12.
- [2] R. M. Nishikawa, R. A. Schmidt, J. Papaioannou, R. Onsis, R. A. Haldemann Heusler, M. L. Giger, D. E. Wolverton, C. Comstock, and K. Doi, "Performance of a prototype clinical intelligent mammography workstation," in *Digital Mammography '96*, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, Eds., Amsterdam, The Netherlands: Elsevier, 1996, pp. 93-96.
- [3] J. Roehrig, T. Doi, A. Hasegawa, B. Hunt, J. Marshall, H. Romsdahl, A. Schneider, R. Sharbaugh, and W. Zhang, "Clinical results with R2 ImageChecker system," in *Digital Mammography*, N. Karssemeijer, M. Thijsen, J. Hendriks, and L. van Erming, Eds., Dordrecht, The Netherlands: Kluwer, 1998, pp. 395-400.
- [4] D. H. Davies and D. R. Dance, "Automatic computer detection of clustered calcifications in digital mammograms," *Phys. Med. and Biol.*, vol. 35, no. 8, pp. 1111-1118, 1990.
- [5] H.-P. Chan, K. Doi, C. J. Vyborny, R. A. Schmidt, C. E. Metz, K. L. Lam, T. Ogura, Y. Wu, and H. MacMahon, "Improvement in radiologists' detection of clustered microcalcifications on mammograms: The potential of computer-aided diagnosis," *Investigative Radiol.*, vol. 25, no. 10, pp. 1102-1110, 1990.
- [6] H. Yoshida, K. Doi, R. M. Nishikawa, M. L. Giger, and R. A. Schmidt, "An improved computer-assisted diagnostic scheme using wavelet transform for detection of clustered microcalcifications in digital mammograms," *Academic Radiol.*, vol. 3, pp. 621-627, 1996.
- [7] F. Lefebvre, H. Benali, R. Gilles, E. Kahn, and R. Di Paola, "A fractal approach to the segmentation of microcalcifications in digital mammograms," *Med. Phys.*, vol. 22, no. 4, pp. 381-390, 1995.
- [8] E. A. Krupinski and C. F. Nodine, "Gaze duration predicts the locations of missed lesions in mammography," in *Digital Mammography*, A. G. Gale, S. M. Astley, D. R. Dance, and A. Y. Cairns, Eds., Amsterdam, The Netherlands: Elsevier, 1994, pp. 399-403.
- [9] E. A. Sickles, "Breast calcifications: Mammographic evaluation," *Radiol.*, vol. 160, no. 2, pp. 289-293, 1986.
- [10] A. H. Tulus, "Mammographic-pathological correlations," in *Radiological Diagnosis of Breast Diseases*, M. Friedrich and E. A. Sickles, Eds., Berlin, Germany: Springer-Verlag, 1997, pp. 13-31.
- [11] N. Karssemeijer, "Adaptive noise equalization and recognition of microcalcification clusters in mammograms," *Int. J. Pattern Recognit. Artificial Intell.*, vol. 7, no. 6, pp. 1357-1376, 1993.
- [12] W. P. Kegelmeyer, Jr. and M. C. Allmen, "Dense feature maps for detection of calcifications," in *Digital Mammography*, A. G. Gale, S. M. Astley, D. R. Dance, and A. Y. Cairns, Eds., Amsterdam, The Netherlands: Elsevier, 1994, pp. 3-12.
- [13] D. Meersman, P. Scheunders, and D. Van Dyck, "Detection of microcalcifications using neural networks," in *Digital Mammography '96*, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, Eds., Amsterdam, The Netherlands: Elsevier, 1996, pp. 287-290.
- [14] M. Kimmig, C. Buitrago-Téllez, H. Gaffner, M. Langer, R. Kober, R. Ossig, and C. Schiekel, "Mikrokalkerkennung im mammographien mittels matched filter," in *Digitale Bildverarbeitung in der Medizin*, B. Arnolds, H. Müller, D. Saupe, and T. Tolksdorf, Eds., Freiburg, Germany: GMDS Universität Freiburg, 1997, pp. 94-97.
- [15] R. N. Suckland and H. I. Hahn, "Wavelet transforms for detecting microcalcifications in mammograms," *IEEE Trans. Med. Imag.*, vol. 15, no. 2, pp. 218-228, 1996.
- [16] M. N. Gürgan, Y. Yardimci, and E. Cetin, "Microcalcification detection using adaptive filtering and Gaussianity tests," in *Proc. 4th Int. Workshop Digital Mammography*, Nijmegen, 1998.
- [17] D. Geman, "Random fields and inverse problems in imaging," in *Lecture Notes in Mathematics 1427*, A. Ancona, D. Geman, and N. Ikeda, Eds., Berlin, Germany: Springer-Verlag, 1990, pp. 116-193.
- [18] G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*. Berlin, Germany: Springer-Verlag, 1995.
- [19] T. Netsch, "A scale-space approach for the detection of clustered microcalcifications in digital mammograms," in *Digital Mammography '96*, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, Eds., Amsterdam, The Netherlands: Elsevier, 1996, pp. 301-306.
- [20] ———, "Automated detection of clustered microcalcifications in digital mammograms," Ph.D. dissertation, Univ. of Bremen, Bremen, Germany, 1998.
- [21] C. E. Metz, "ROC methodology in radiologic imaging," *Investigative Radiol.*, vol. 21, no. 9, pp. 720-733, 1986.
- [22] D. P. Chakraborty and L. H. Winter, "Free response methodology: Alternate analysis and a new observer-performance experiment," *Radiol.*, vol. 174, no. 3, pp. 873-881, 1990.
- [23] D. Blostein and N. Abuja, "A multiscale region detector," *Comput. Vision, Graph., Image Processing*, vol. 45, pp. 22-41, 1989.
- [24] T. Lindeberg, Ed., *Scale-Space Theory in Computer Vision*. Boston, MA: Kluwer, 1994.

- [25] B. M. ter Haar Romeny, Ed., *Geometry-Driven Diffusion in Computer Vision*. Dordrecht, The Netherlands: Kluwer, 1994.
- [26] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.
- [27] F. Neyenssac, "Contrast enhancement using the Laplacian-of-a-Gaussian filter," *CVGIP: Graph. Models Image Processing*, vol. 55, no. 6, pp. 447-463, 1993.
- [28] A. P. Witkin, "Scale-space filtering," in *Proc. Int. Joint Conf. Artificial Intelligence*, Karlsruhe, Germany, 1983, pp. 1019-1026.
- [29] A. L. Yuille and T. A. Poggio, "Scaling theorems for zero crossings," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 15-25, Jan. 1986.
- [30] C. J. Everts, K. Berkner, and W. Berghorn, "A local multiscale characterization of edges applying the wavelet transform," in *Proc. Nato A.S.I. Fractal Image Encoding Analysis*, July 1995.
- [31] H.-P. Chan, K. Doi, S. Galbraith, C. J. Vyborny, H. MacMahon, and P. M. Jokich, "Image feature analysis and computer-aided diagnosis in digital radiography: Automated detection of microcalcifications in mammography," *Med. Phys.*, vol. 14, no. 4, pp. 538-548, 1987.
- [32] N. Karssemeijer and L. van Erning, "Iso-precision scaling of digitized mammograms to facilitate image analysis," *SPIE Proc.*, vol. 1445 pp. 166-177, 1991.
- [33] T. Ena, K. Doi, R. M. Nishikawa, and J. Papaioannou, "Image feature analysis and computer-aided diagnosis in mammography: Reduction of false-positive clustered microcalcifications using local edge-gradient analysis," *Med. Phys.*, vol. 22, no. 2, pp. 161-169, 1995.
- [34] A. Rosenfeld and A. C. Kak, *Digital Picture Processing Volume 2*. San Diego, CA: Academic, 1982.
- [35] C. E. Priebe, "Detection of microcalcifications clusters in digital mammography via the spatial scan statistic with stochastic scan partitions," in *Proc. 4th Int. Workshop Digital Mammography*, Nijmegen, 1998.
- [36] R. M. Nishikawa, M. L. Giger, K. Doi, C. J. Vyborny, and R. A. Schmidt, "Computer-aided detection of clustered microcalcifications: An improved method for grouping detected signals," *Med. Phys.*, vol. 20, no. 6, pp. 1661-1666, 1993.
- [37] H.-P. Chan, L. T. Niklason, D. M. Ikeda, K. L. Lam, and D. D. Adler, "Digitization requirements in mammography: Effects on computer-aided detection of microcalcifications," *Med. Phys.*, vol. 21, no. 7, pp. 1203-1211, 1994.

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**